Face Recognition Vendor Test (FRVT)
(Ongoing)

FRVT 1:1 Verification
Frequently Asked Questions

Version 0.2

Patrick Grother and Mei Ngan

Contact via **frvt@nist.gov**

Image Group
Information Access Division
Information Technology Laboratory

National Institute of Standards and Technology
U.S. Department of Commerce

January 30, 2017

Table 1 – Questions Related to images and subjects

#	Question	Answer
1.	You enumerate certain image types so that algorithms can alter their processing, will NIST provide examples of these?	In short, no. The properties of visa and passport photos are widely known because of their (approximate) conformance to ISO/IEC 19794-5:2005, so NIST will not provide those. A sample of mugshots is available in the MEDS[1] dataset. A sample of photojournalism images is available in the IJB-A[2] set, and perhaps LFW[3]. Wild images can be obtained in volume from social media e.g. flickr. We consider wild images to include photos from both amateur and professional photographers. We cannot provide child-exploitation images, and we advise developers that access and possession of such images is illegal in most jurisdictions.
2.	So why does NIST not provide training images?	NIST does not provide training data, because a) we cannot, and b) because that is commercial practice. Face recognition is typically deployed using "off-the-shelf" software that is not re-trained on the end-user data. This is opposite to the case in academic research which very often includes training data.
3.	What range of subject ages are present in NIST image collections.	We will evaluate on images from children and adults, so the age range varies from a few months old to at least 90 years old.
4.	Will a minimum and maximum face size be provided?	No. But note that if the image type label is "wild", "photojournalism", "unknown" or "child exploitation" then very widely divergent sizes are typical. For "mugshot" and "iso" types, then we intend to provide faces with interocular distance 40 <= IOD <= 300 pixels.
5.	What are the main characteristics of faces in child exploitation imagery?	The problems are five-fold: 1. Faces are from children aged anywhere from 0 to 17 inclusive. 2. The faces can have very high roll angles, consistent with a subject lying down, in addition to unconstrained yaw and pitch. 3. The faces can be occluded by hair, limbs and other objects. 4. The faces can have very low to very high resolution. Some photos are similar to generic "Wild" photos.
6.	Will NIST run tests on 3D images.	We do not anticipate any tests with 3D images.
7.	Will NIST compare images of two types, e.g. Wild with VISA.	Yes, because it is reflective of some operational use-cases, e.g. Passport e-Gates. NIST will additionally compare images within type e.g. mugshot vs. mugshot.

[1] NIST Special Database 32 – Multiple Encounter Dataset (MEDS) - https://www.nist.gov/itl/iad/image-group/special-database-32-multiple-encounter-dataset-meds
[2] IARPA Janus Benchmark A (IJB-A) - https://www.nist.gov/programs-projects/face-challenges
[3] Labeled Faces in the Wild (LFW) - http://vis-www.cs.umass.edu/lfw/

Table 2 – Questions Related to Number and Schedule of Algorithm Submissions

#	Question	Answer
8.	How many algorithms can we submit?	For an organization that has never participated in FRVT Ongoing, two algorithms may be submitted. For an organization that has previously participated, see next questions.
9.	How many algorithms do you recommend we submit?	We recommend you submit one algorithm now. Then consider the initial test reports that NIST releases. NIST will run on limited data initially, adding new datasets, and new analyses over time. Once the test reporting is stabilized, a new developer might submit two algorithms to NIST.
10.	How often may developers submit to the test?	Developers may submit one or two initial algorithms. Once two algorithms have been evaluated, a new algorithm may be submitted no earlier than three calendar months since the most recent submission. So, for example, if Developer X submitted X1 on March 3, and submitted X2 on April 1, she may submit a new algorithm X3 on July 1 or anytime later. X3 will replace X1 in the main results listings. An X4 variant could be submitted October 1.
11.	Why does NIST limit the frequency of submissions	To limit the rate at which developers can tune to the test data. To limit the computational resources NIST requires to run the tests.

Table 3 – Questions Related to Reporting of Results

#	Question	Answer
12.	Where will NIST report results?	NIST will place PDF test reports on the FRVT website, and include HTML tables there also. NIST may include results in NIST interagency reports, academic publications, and in conference, workshop and seminar presentations.
13.	NIST will produce a test report for an algorithm, what other algorithms will appear in the test report.	NIST will regenerate and republish test reports as soon as any new algorithm completes a test. So, for example, the report for a university's X1 algorithm will be updated a year later when results from a company's Q4 algorithm are available. This is intended to support comparison of algorithms.
14.	How many algorithms will NIST list on the main FRVT website?	Two. NIST will designate algorithms as new or old. NIST will treat algorithm submissions as a FIFO queue. The two most recent algorithms will be regarded as "new". NIST will run the new algorithms on all currently active tests. When an updated algorithm is submitted, the oldest algorithm will be assigned "old" status, and may not be run on new test datasets. So, for example, suppose a year from now NIST obtains a new database of passport photographs, NIST will run the two most recent ("new") algorithms on it, but will not normally run the "old" algorithms. It is possible in this scenario for an old algorithm to outperform the more recent algorithms and not appear on the main results pages.
15.	Will NIST report results for old algorithms?	NIST will archive all test reports on the FRVT website. NIST will not ordinarily list results on the main page, but reserves the right to do so.
16.	What performance metrics will NIST report, beyond accuracy and speed.	NIST will report matching error rates (FNMR, FMR), for whole datasets, for specific populations groups, and for specific types of images. We will also report processing durations, software and data sizes, template sizes, memory consumption. We reserve the right to report other metrics, for example on the biometric zoo. We may report any anomalous behavior.

Table 4 – Questions Related to Software Properties and Behavior

#	Question	Answer
17.	What is a submission to FRVT?	A submission is software implementing the API and other requirements given in the FRVT *Still Face 1:1 Verification Concept, Evaluation Plan and API.* The document is linked from here: https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt-ongoing
18.	What is the order of NIST activity?	1. Validation. 2. Timing. 3. Full evaluation. 4. Report generation. 5. Publication. Completion of each step is a necessary condition for progression to the next.
19.	What is the point of validation?	Validation serves solely to ensure software run at NIST gives identical outputs to software run by the developer. The validation images are not necessarily representative of images NIST will use. The images do not support training, nor testing.
20.	What happens if we fail validation?	Usually validation succeeds: Both templates and scores are identical. If validation fails, we will email you with details of the observed differences. You may then modify and re-submit the algorithm, or you may request that NIST continues to test the algorithm ignoring the failed validation.
21.	Will NIST evaluate face recognition software running on Windows or MacOS?	No. We run on a cluster of computers equipped with a linux variant. This is economical and reflective of many government and commercial back-ends. We anticipate the organizations should be able to separate their core algorithmic code from, from example access control or GUI-driven products.
22.	What happens if the software is too slow?	NIST will send email with indicating the speed and ask you to resubmit a faster implementation. NIST will discontinue execution. Results will not be posted. We reserve the right to post that software was received and was unable to complete the test.
23.	What happens if software crashes?	NIST will send email with whatever diagnostic information we have, and will discontinue evaluation. We will ask you to resubmit a fixed implementation. If the revised software fails, we will repeat this process. However, if we judge that the software is too unreliable, we will ask you to do a more extensive software review and ask you to submit three calendar months later.
24.	What happens if software does not make a template?	NIST will send email with whatever diagnostic information we have, and will discontinue evaluation. We will ask you to resubmit a fixed implementation. NOTE: Given an input MULTIFACE, the createTemplate() function must return a template even if the software cannot detect a face or chooses not to extract features. The function may return a non-zero error code as enumerated in the API document, but must *always* produce a template. See the documentation for createTemplate() in the API document.
25.	What libraries can we supply, e.g. openCV or caffe?	You may provide zero or more supporting libraries. These libraries should all reside in one lib directory with the main FRVT library that implements our API. The standard libraries can be seen in our linux distribution. Any other libraries should be sent to us.
26.	Can our software contact servers?	No, your code must not attempt to access any device or resource except main memory, unless we explicitly specify it in the API. If NIST detects that software attempts to connect to, or open, or read, any server, socket, connection, port or other resource, we will stop all tests for an organization, delete all test reports. We may disclose the violation and take additional steps.